A Comparison of Active Classification Methods for Content-Based Image Retrieval

Philippe-Henri Gosselin and Matthieu Cord

ETIS CNRS UMR 8051

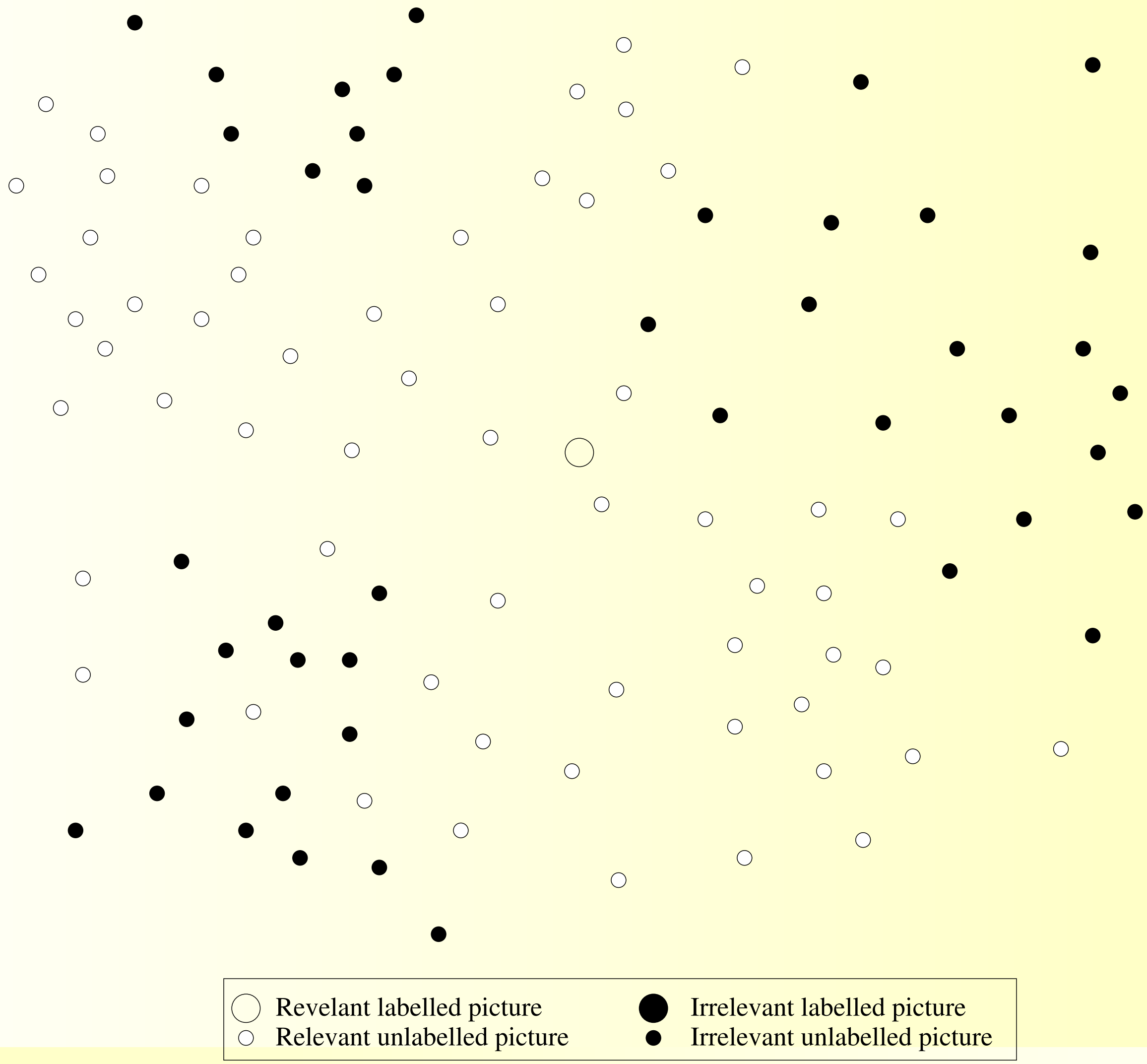
6, av du Ponceau 95014 CERGY-PONTOISE Cedex

gosselin@ensea.fr

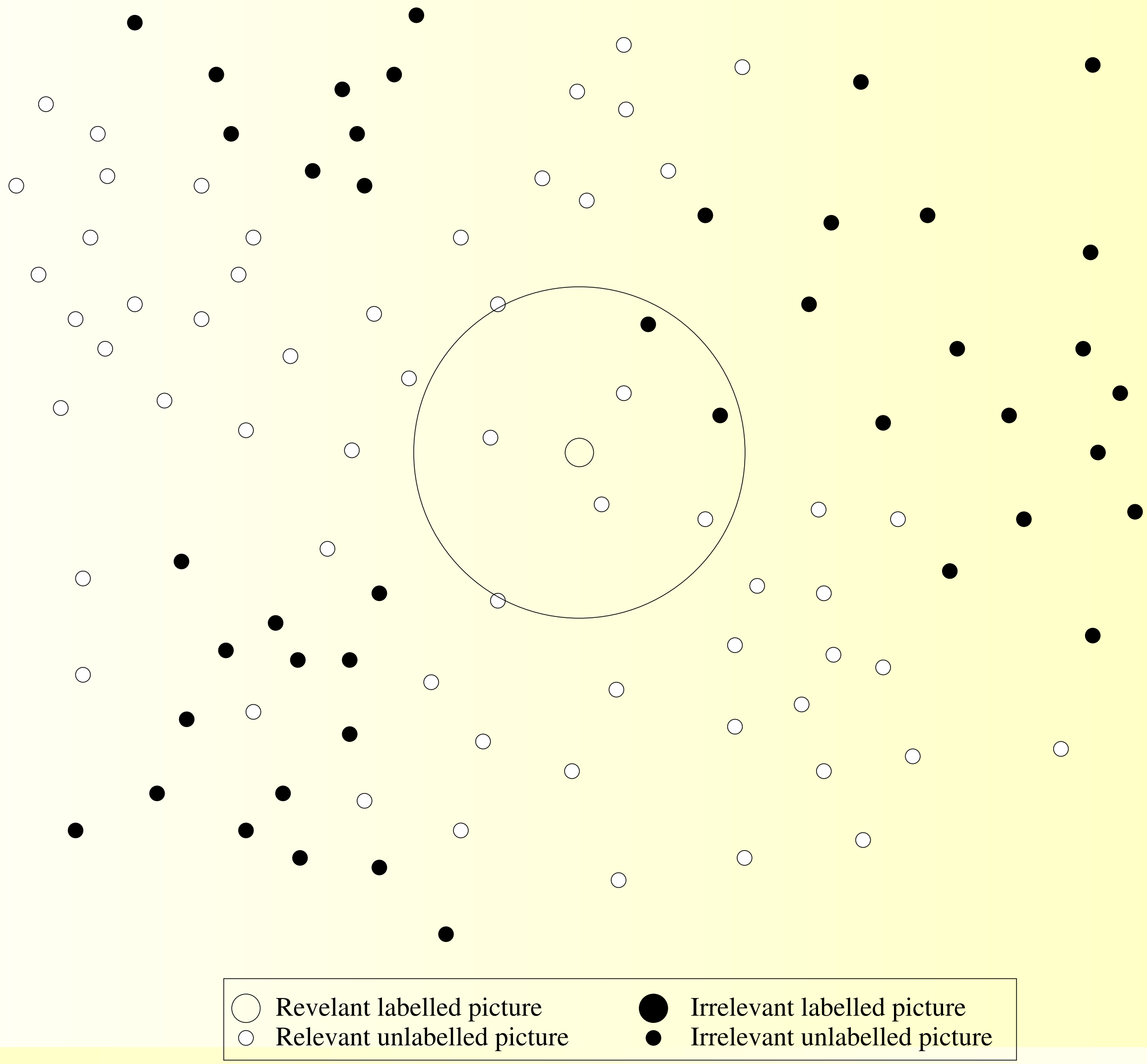
Content-Based Image Retrieval

- Retrieve large categories of pictures in generalist image database;
- Vector-based description of images;
- Statistical learning approach;
- Relevance feedback.

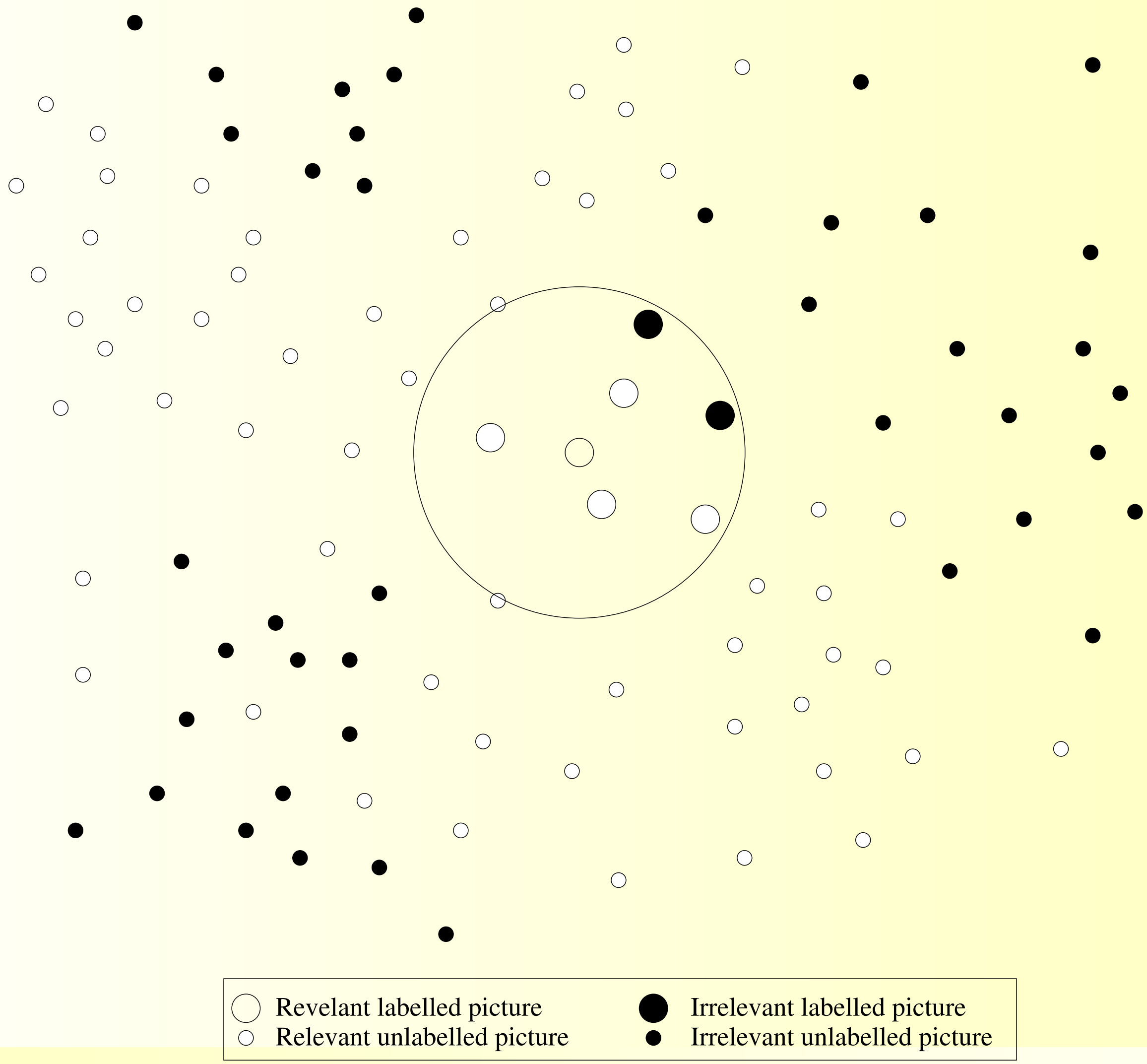
Active binary classification



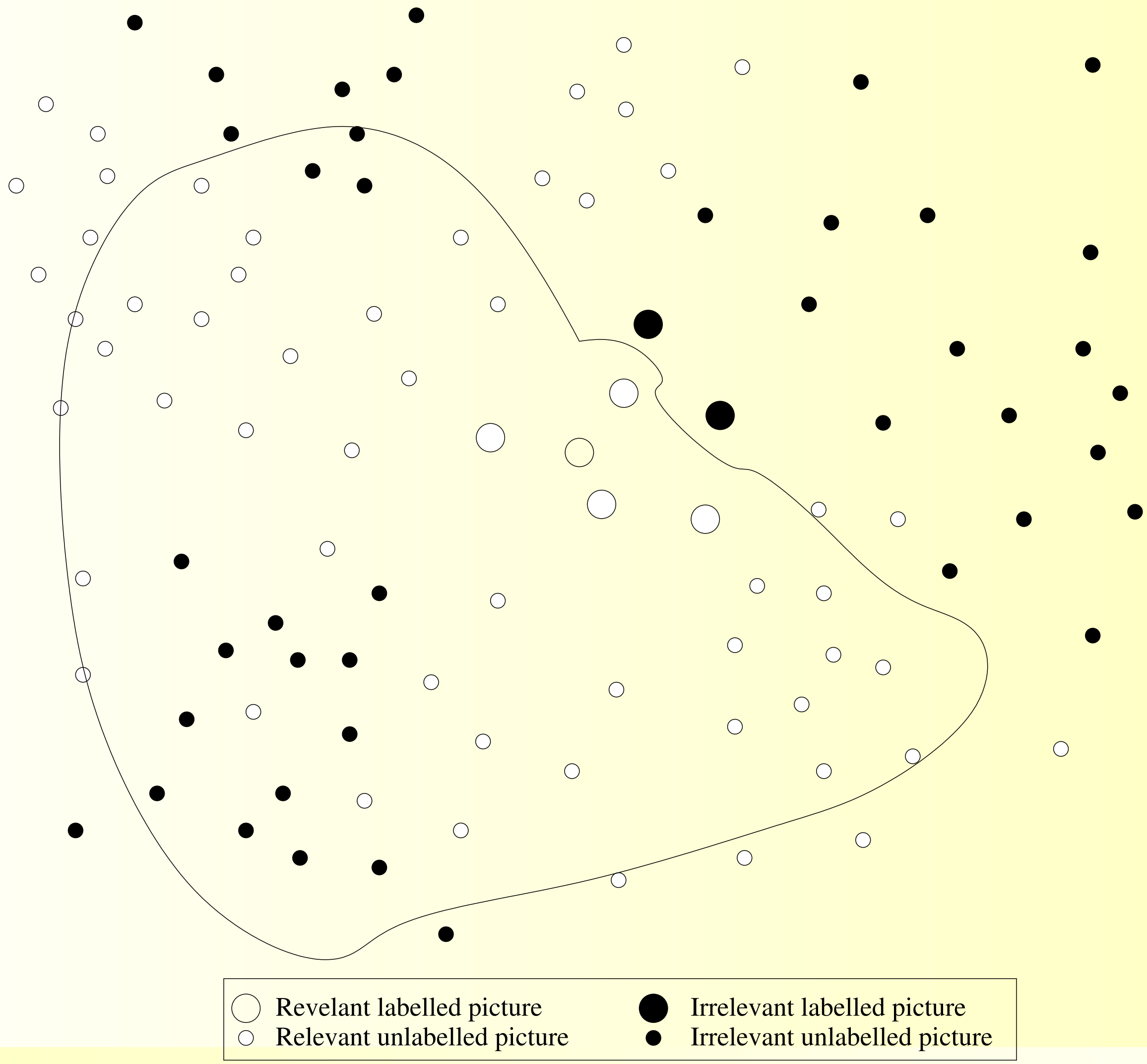
Active binary classification



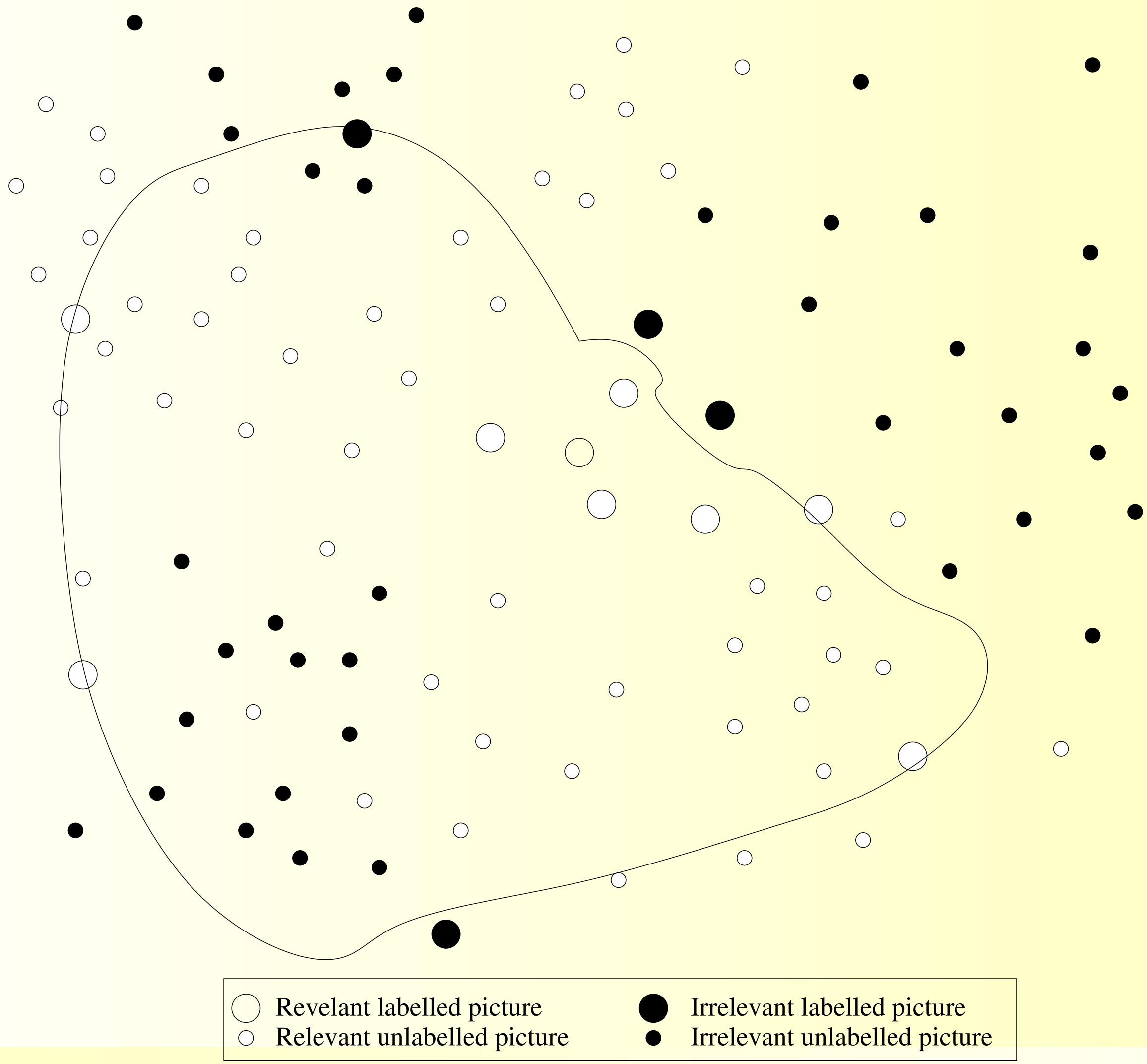
Active binary classification



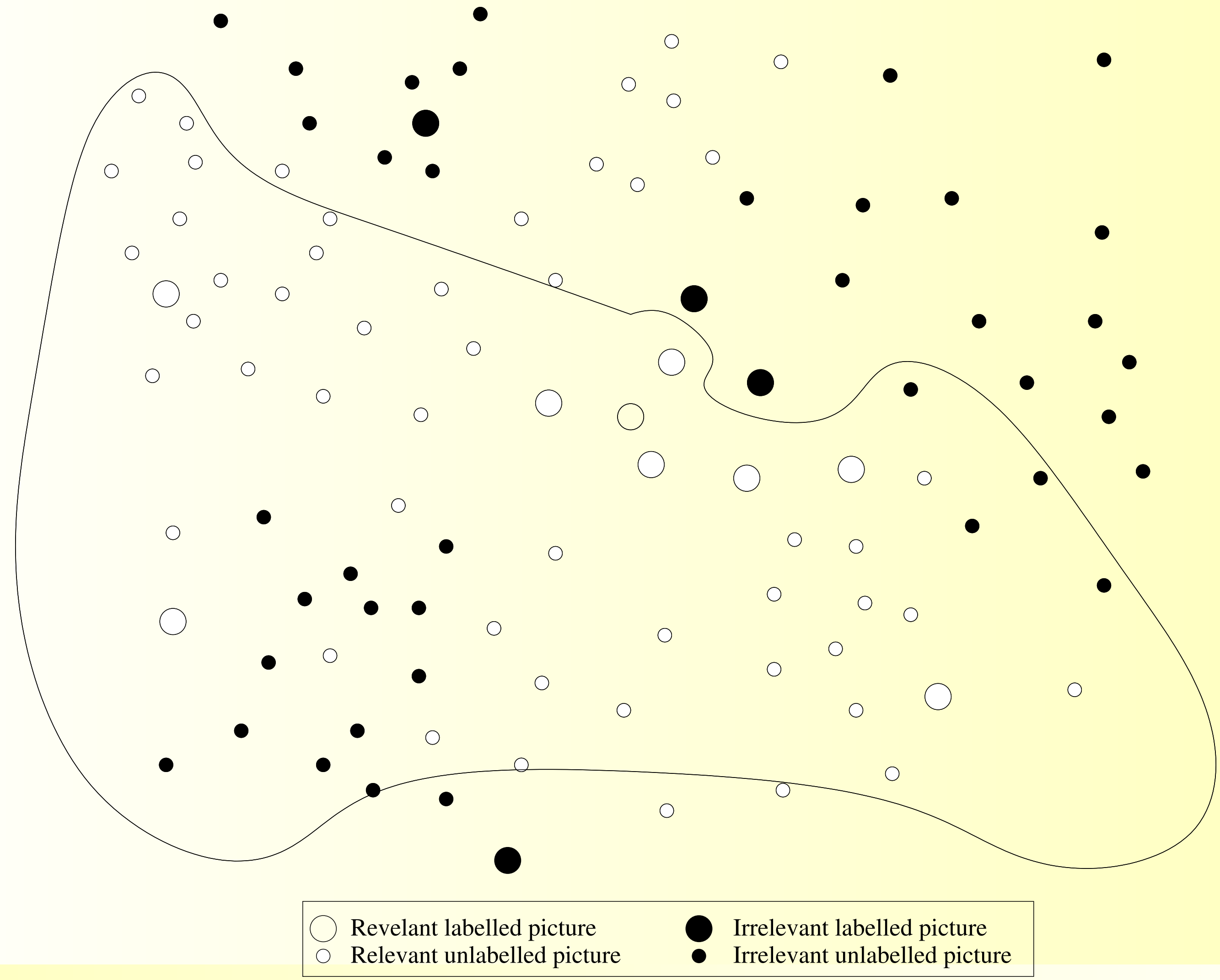
Active binary classification



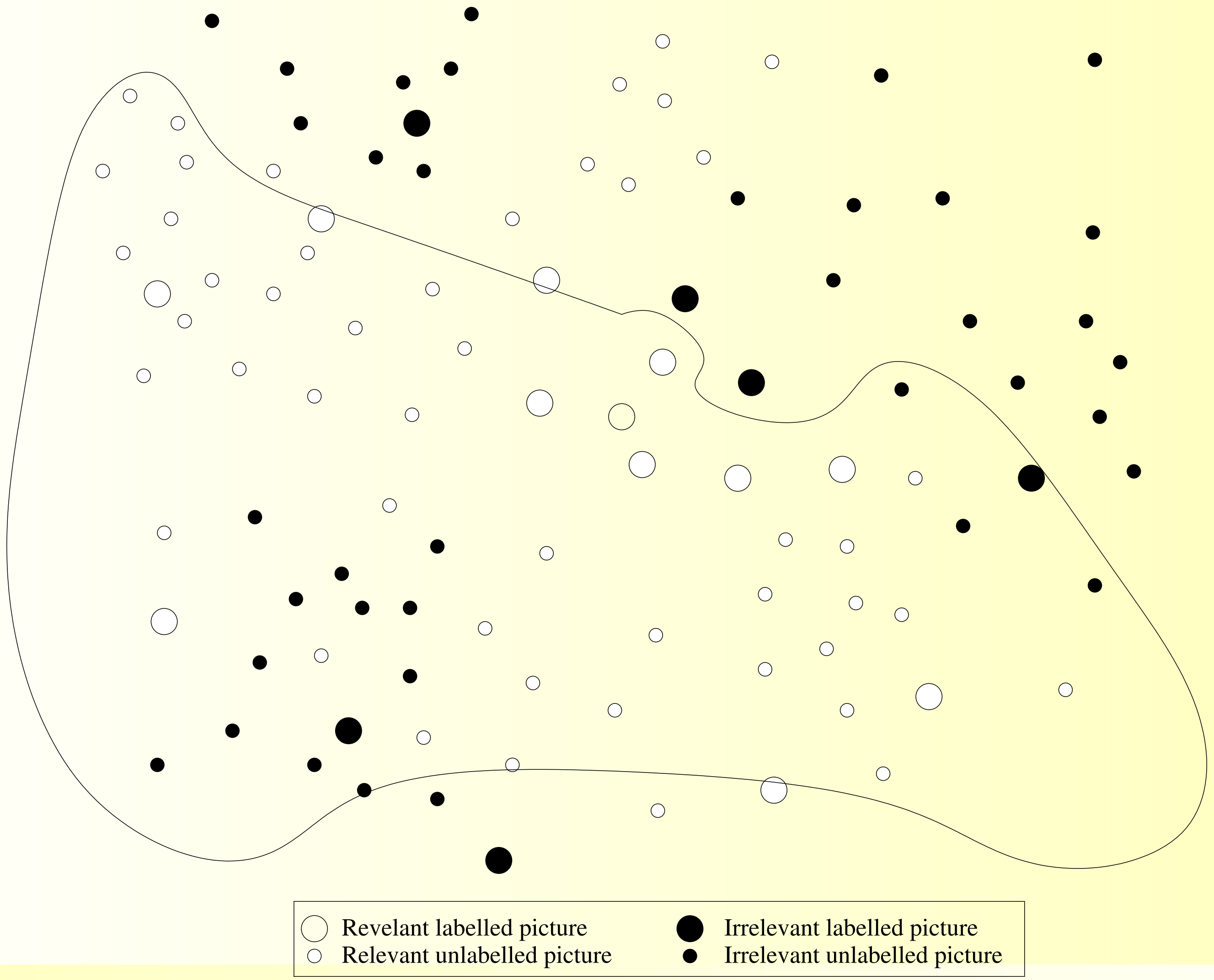
Active binary classification



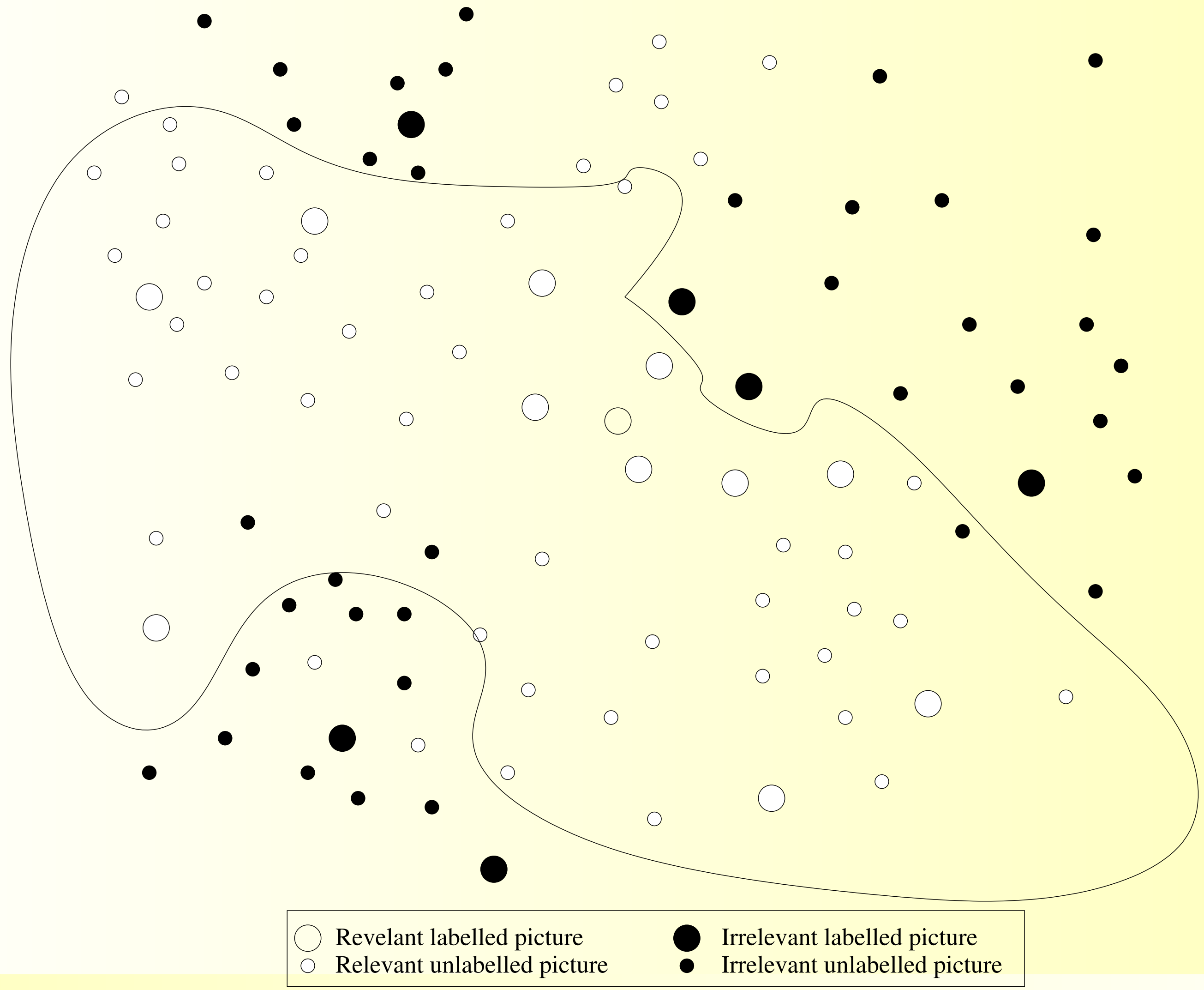
Active binary classification



Active binary classification



Active binary classification



Specific characteristics

(c1) High dimension and non-linearity of input space;

(c2) Few training data;

(c3) Many unlabelled data;

(c4) Active learning;

(c5) Unbalanced training data.

Outline

- Introduction to Classification Methods;
- Does Transduction improve Classification ?;
- Does Active learning improve Classification ?;
- Experiments.

Classification

Classification methods

Three well-known methods:

- Bayes Classifiers;
- k-Nearest Neighbors;
- Support Vector Machines.

These methods work well with low dimension and linear input space, but not with high dimension and non-linear input space (c1):

- $\longrightarrow$ Use of a kernel function to induce a feature space;
- $\longrightarrow$ The induced feature space replaces the input space, but with nice properties.

"Kernelization"

Kernelization of SVM:

- SVM decision function:

$$f(\mathbf{x}) = \sum_{i=1}^{N} y_i \alpha_i^{\star} < \mathbf{x}, \mathbf{x}_i > +b \qquad (1)$$

- "Kernelized" version:

$$f(\mathbf{x}) = \sum_{i=1}^{N} y_i \alpha_i^{\star} K(\mathbf{x}, \mathbf{x}_i) + b \qquad (2)$$

When a method cannot be directly "kernelized": KPCA.

Kernels

- Usual kernels: Linear, Polynomial, Sigmoid, RBF, ...
- Choice of a kernel depends on the database and its usage:

$\rightarrow$ Different levels of performances for two different kernels;

- In our experiments: Gaussian kernels give the best results
  - $\rightarrow$ The most adapted to CBIR;
  - $\rightarrow$ In the following experiments: Gaussian kernels with χ^2 distance, because feature vector are distributions.

Does Transduction improve classification ?

Transductive methods

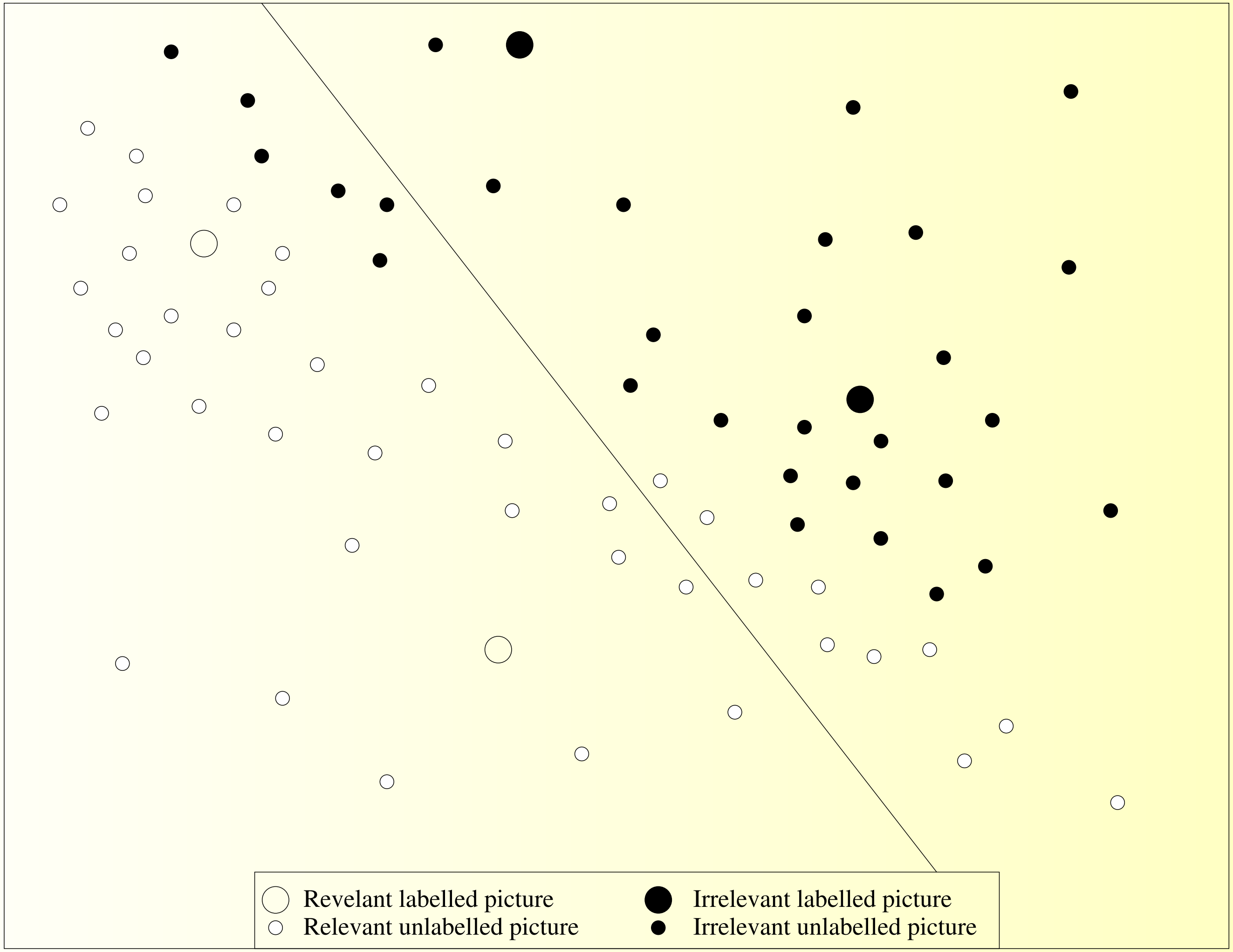
Deal with the two following characteristics:

(c2) Few training data;

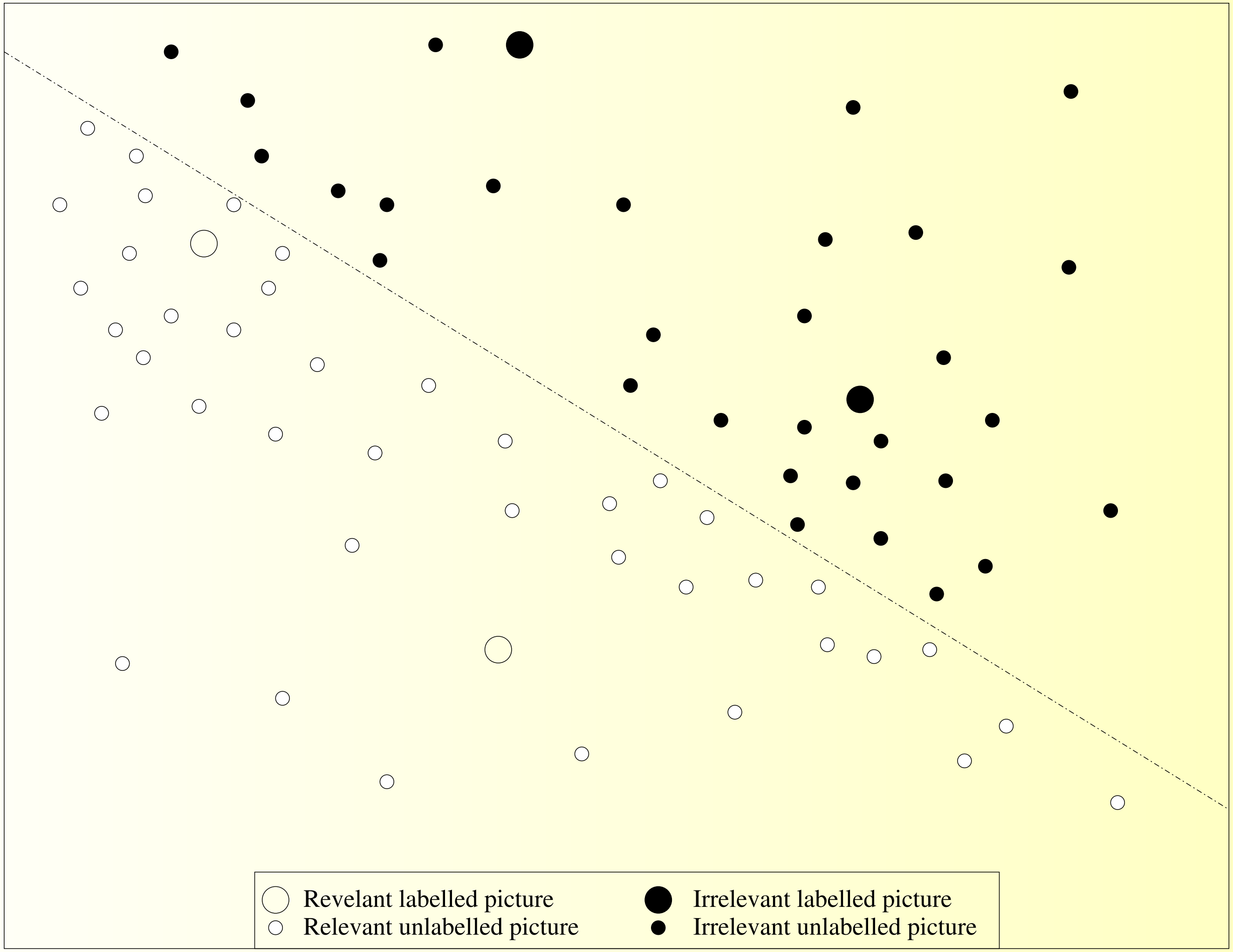
(c3) Many unlabelled data.

Transductive methods use unlabelled data to compensate scarcity of training data.

Inductive SVM



Transductive SVM (Joachims)



Experiments

- Very high computational cost compared to induction:
  → In the following experiments: 1 minute for Transduction, 2 seconds for Induction;
- The tuning of the size of relevant class is not trivial (c5);
- Performances are sometimes increased, sometimes decreased
  → Possible cause: Structural hypothesis made by this transducer does not correlate with database structure.
- Actually: Transduction does not seems to be adapted to CBIR.

Does Active Learning improve classification ?

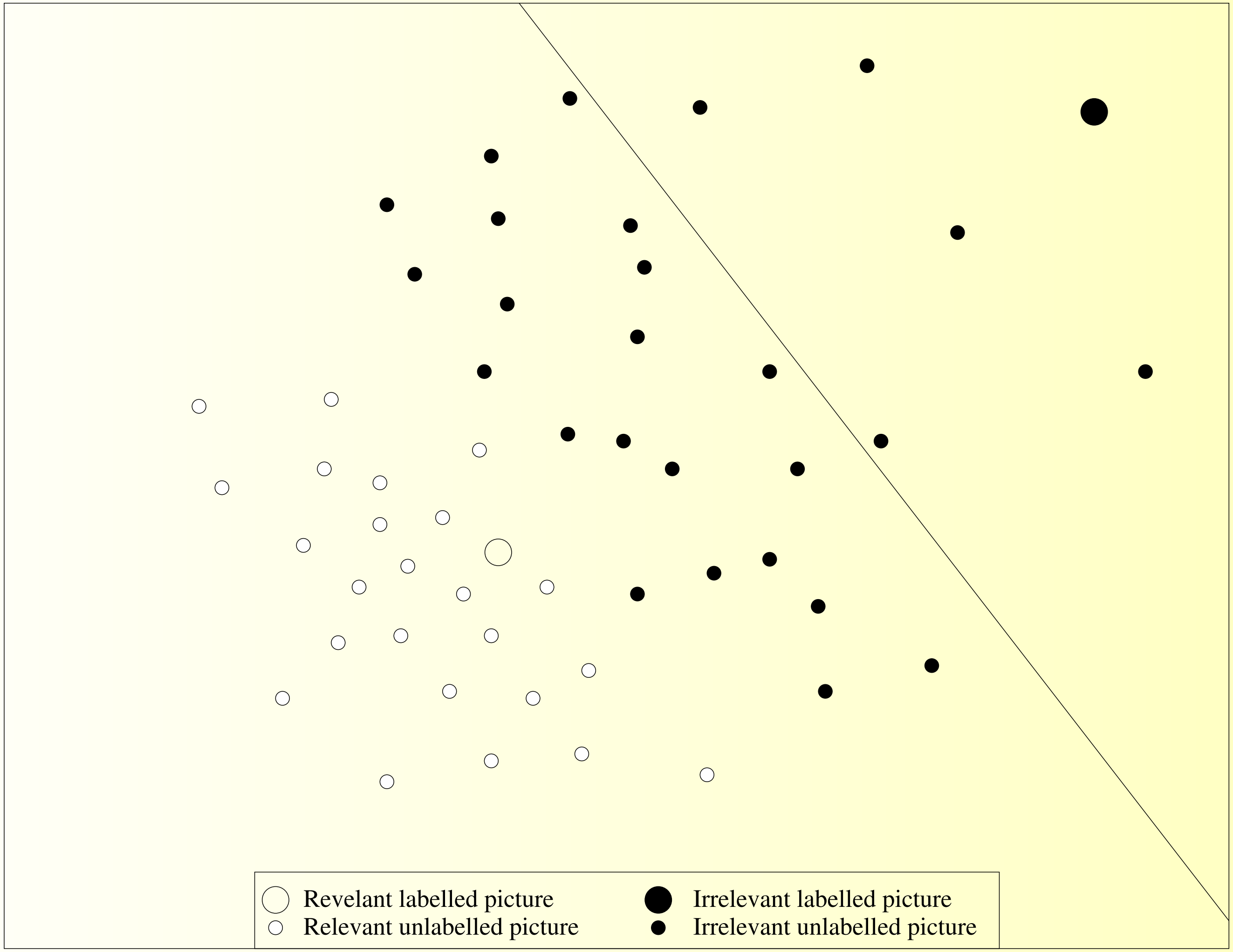
Active Learning

Deal with the following characteristic:

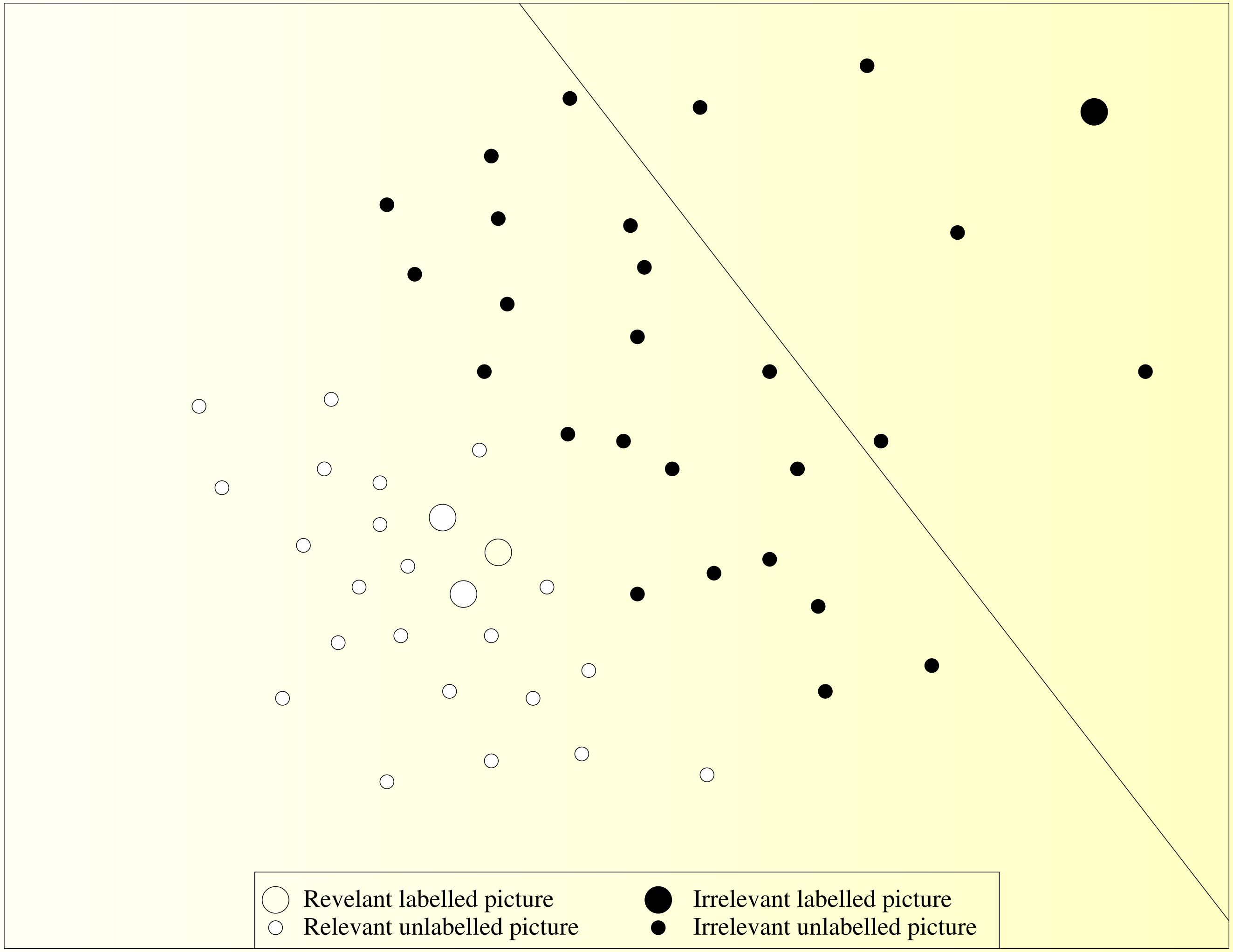
(c2) Few training data;

Active learning methods optimize training data to compensate its scarcity.

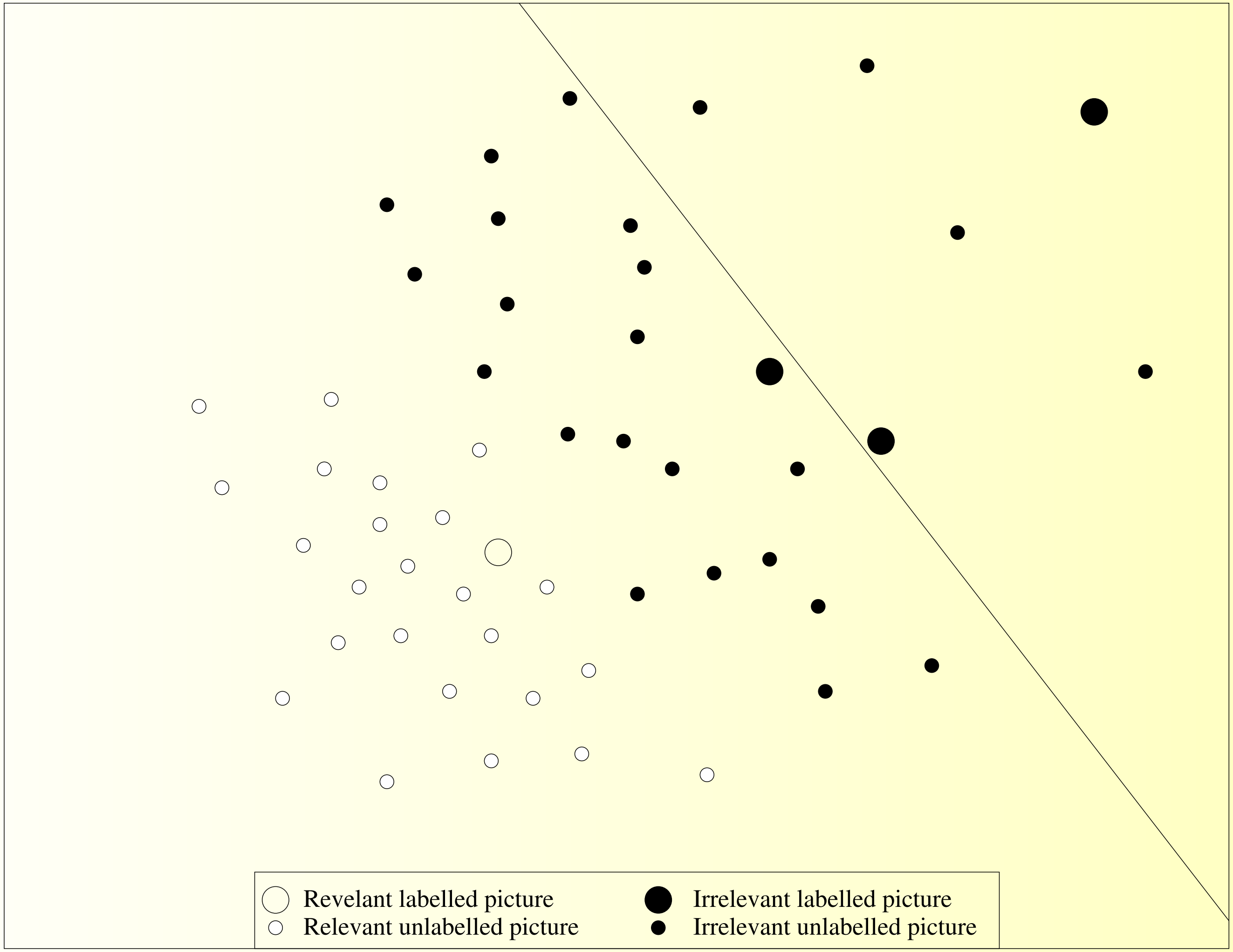
Initial training data



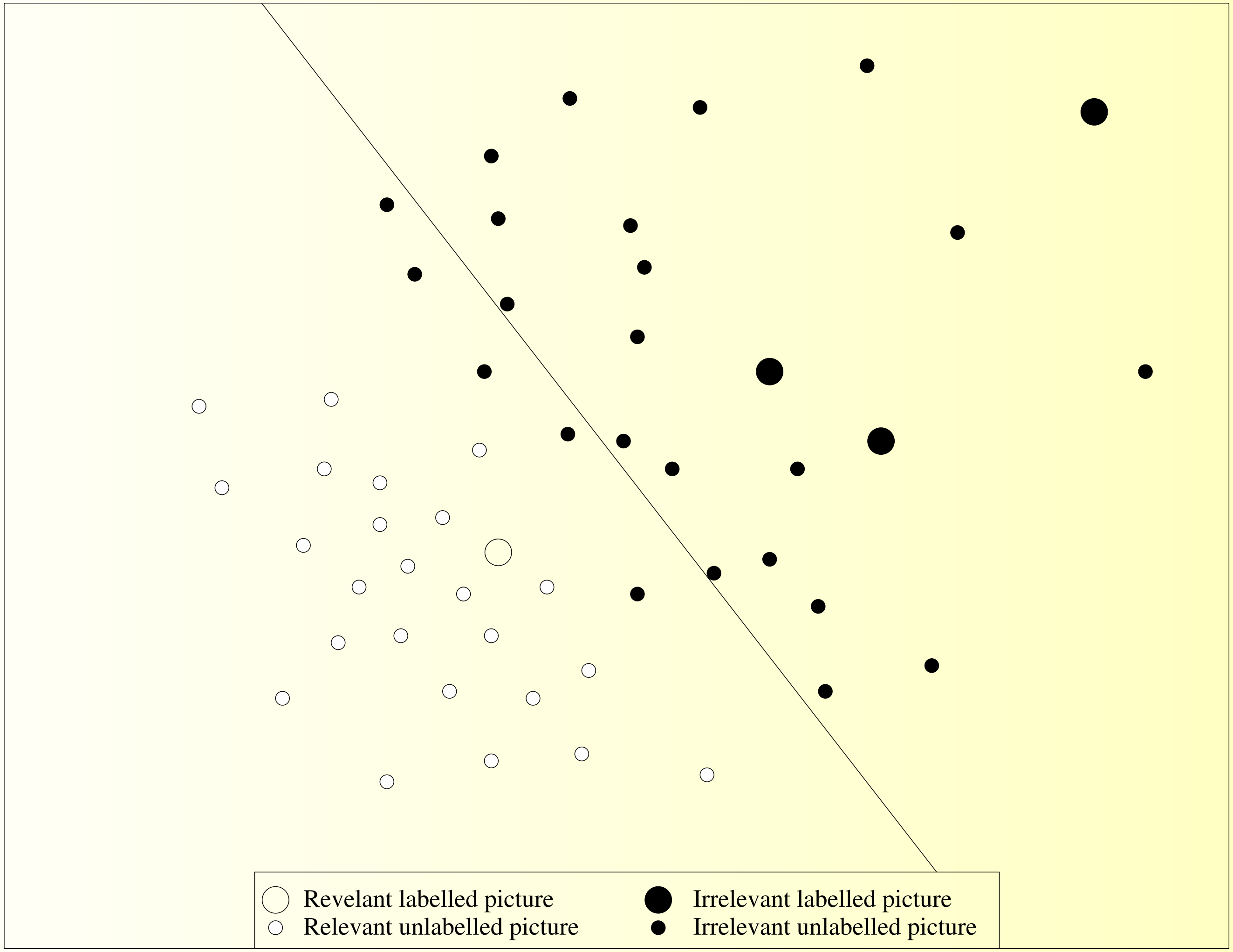
Basic active learning



Pictures the most difficult to classify



Classification is enhanced



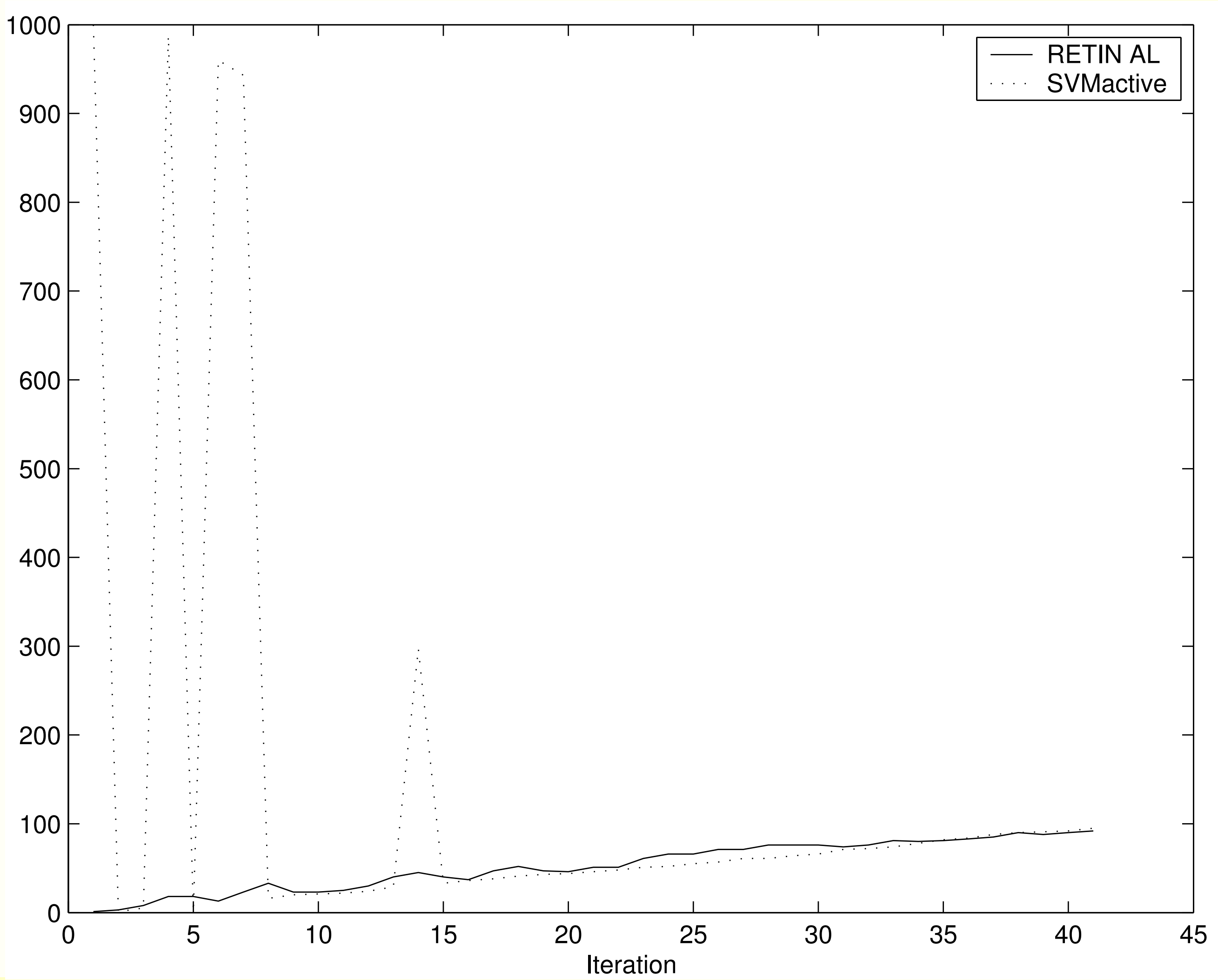
SVM$_{active}$ (Tong)

- Method with strong theoretical foundation;
- The closer to the SVM boundary an image is, the less reliable its classification is:
  - $\rightarrow$ Choose images closest to the SVM boundary;
- Important assumption: a reliable estimation of the boundary:
  - $\rightarrow$ Few training data (c2) and unbalanced training data (c5) characteristics do not allow such assumption;
  - $\rightarrow$ Start with "enough" example.

Proposed Method

- Keep SVM$_{active}$ principle;
- Do not use an estimate to the boundary;
- First idea: have as many relevant as irrelevant labels:
  - $\longrightarrow$ System proposes new labelling further from relevant class center when labels are almost relevant;
  - $\longrightarrow$ And vice-versa;
- Second idea: increase sparseness of the training data:
  - $\longrightarrow$ System computes a clustering around the boundary;
  - $\longrightarrow$ and chooses pictures in different clusters.

Distance to center ac. to feedback steps



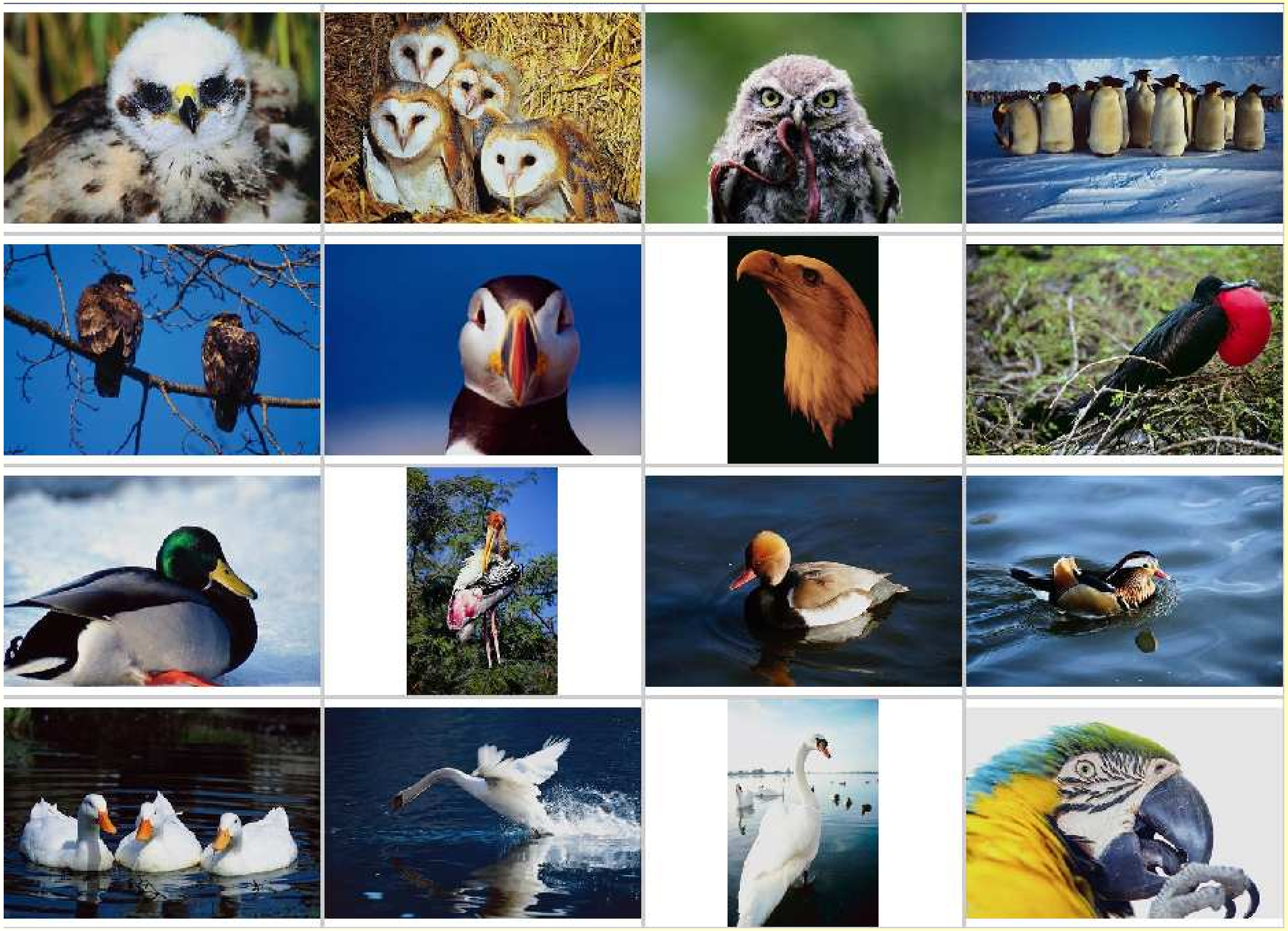
Experiments

Image Database

- Extract of 6,000 pictures from COREL photo database;
- Feature distributions: $L^\star a^\star b^\star$ colors and Gabor filters;
- 11 categories:
  - → From 100 to 600 images;
  - → From simple (monomodal) to very complex (multimodal), according to features;
- Performance metric: average precision (sum of precision/recall curve).

Simple category: Doors

Complex category: Birds



Comparison of 7 active classifiers

Classifier Active learner	Bayes Basic	kNN Basic	SVM Basic	Bayes Proposed	kNN Proposed	SVM Proposed	SVM active
birds	19	29	31	20	29	31	31
castles	15	17	38	17	18	38	38
caverns	72	75	77	73	75	78	75
dogs	22	28	58	21	32	58	58
doors	86	88	89	91	90	93	83
Europe	26	30	33	26	30	35	35
flowers	56	59	60	64	63	67	57
food	58	62	66	64	66	71	59
mountains	30	42	54	39	39	54	54
objects	60	69	75	67	67	78	76
savannah	56	58	62	60	59	68	56

Different Complexities

Classifier Active learner	Bayes Basic	kNN Basic	SVM Basic	Bayes Proposed	kNN Proposed	SVM Proposed	SVM active
birds	19	29	31	20	29	31	31
castles	15	17	38	17	18	38	38
caverns	72	75	77	73	75	78	75
dogs	22	28	58	21	32	58	58
doors	86	88	89	91	90	93	83
Europe	26	30	33	26	30	35	35
flowers	56	59	60	64	63	67	57
food	58	62	66	64	66	71	59
mountains	30	42	54	39	39	54	54
objects	60	69	75	67	67	78	76
savannah	56	58	62	60	59	68	56

SVM is the best classifier

Classifier Active learner	Bayes Basic	kNN Basic	SVM Basic	Bayes Proposed	kNN Proposed	SVM Proposed	SVM active
birds	19	29	31	20	29	31	31
castles	15	17	38	17	18	38	38
caverns	72	75	77	73	75	78	75
dogs	22	28	58	21	32	58	58
doors	86	88	89	91	90	93	83
Europe	26	30	33	26	30	35	35
flowers	56	59	60	64	63	67	57
food	58	62	66	64	66	71	59
mountains	30	42	54	39	39	54	54
objects	60	69	75	67	67	78	76
savannah	56	58	62	60	59	68	56

Proposed a.l. has the best results

Classifier Active learner	Bayes Basic	kNN Basic	SVM Basic	Bayes Proposed	kNN Proposed	SVM Proposed	SVM active
birds	19	29	31	20	29	31	31
castles	15	17	38	17	18	38	38
caverns	72	75	77	73	75	78	75
dogs	22	28	58	21	32	58	58
doors	86	88	89	91	90	93	83
Europe	26	30	33	26	30	35	35
flowers	56	59	60	64	63	67	57
food	58	62	66	64	66	71	59
mountains	30	42	54	39	39	54	54
objects	60	69	75	67	67	78	76
savannah	56	58	62	60	59	68	56

Conclusion

Conclusion

(c1) High dimension and non-linearity of input space;
$\rightarrow$ Use a kernel function;

(c2) Few training data;
$\rightarrow$ Use an Active learner;

(c3) Many unlabelled data;
$\rightarrow$ Transduction does not seem to be adapted;

(c4) Active learning;
$\rightarrow$ Efficient proposed method;

(c5) Unbalanced training data.
$\rightarrow$ Not discussed.

Future work

- Active learning model for CBIR;
- Application and ground-truth performance evaluation to an artwork database
  - $\rightarrow$ Actually current model is efficient for a 20,000 image database;
- Scalability of model
  - $\rightarrow$ Exploration step.